Impact of visual contact on vocal interaction dynamics of pair bonded birds

Emilie C. Perez, Marie S.A. Fernandez, Simon C. Griffith, Clémentine Vignal, Hedi A. Soula

Impact of visual contact on vocal interaction dynamics of pair bonded birds

Perez EC *[1,2], Fernandez MSA*[1,3], Griffith SC[2], Vignal C[1], Soula HA[3,4]

1. Université de Lyon/Saint-Etienne, Neuro-PSI/ENES CNRS UMR 9197, France

2. Department of Biological Sciences, Macquarie University, SYDNEY, NSW 2109, Australia

3. EPI BEAGLE INRIA F69621 VILLEURBANNE

4. INSERM U1060 INSA LYON F69621 VILLEURBANNE

*: These authors contributed equally to this work

Corresponding author: MSA Fernandez, Université de Lyon/Saint-Etienne, Neuro-PSI/ENES CNRS UMR 9197, France.

E-mail address: marie.fernandez@inria.fr, tel: 33 (0)4 77 48 15 21, fax: 33 (0)4 77 48 51 16

Correspondance:

EC Perez: perezemilie42@yahoo.fr

SC Griffith: simon.griffith@mq.edu.au

C Vignal: Clementine.Vignal@univ-st-etienne.fr

HA Soula: hedi.soula@insa-lyon.fr

Word count: 11472

Animal social interactions usually revolve around several sensory modalities. For birds, these are primarily visual and acoustic. However, some habitat specificities or large distances may temporarily hinder or limit visual information transmission making acoustic transmission a central channel of communication even during complex social behaviours. Here we show the impact of visual limitation on the vocal dynamics between zebra finches partners. Pairs were acoustically recorded during a separation and reunion protocol with gradually decreasing distance without visual contact. Without visual contact, pairs display more correlated vocal exchanges than with visual contact. We also analysed the turn-taking sequences of individuals' vocalisations during an exchange with or without visual contact. We show that in the absence of visual contact, the identity of a vocalising individual is well predicted by the knowledge of the identity of the previous vocaliser. This property is characteristic of a stochastic process called a Markov chain and we show here that deprived of visual contact, turn-taking sequences are Markovian. Thus, both the temporal correlation between the calls of the two partners and Markov properties of acoustic interactions indicate that in the absence of visual clues the decision to emit a call is taken on a very short-term basis and solely on acoustic information (both temporal and identity of caller). Strikingly, when individuals are in visual contact both these features of their acoustic social interactions disappear indicating that birds adapt their calling dynamics to cope with limited visual cues.

Keywords: Markov chains, pair-bond, turn-taking, visual contact, vocal communication, zebra finch

Whilst individual traits usually drive the probability of survival and breeding in a given environment, properties emerging from interactions between mates can also influence the success of a pair, overriding the influence of intrinsic individual quality (Ens, Safriel, & Harris, 1993; Ryan & Altmann, 2001). Many long-term monogamous species of birds show an increase in breeding success with pair bond duration, which is attributed to the improvement in partners' coordination over time (mate familiarity effect, Black, 2001; Black & Hulme, 1996; Forslund & Pärt, 1995). The strength of coordination and synchronization of behaviours within a pair may at least partly depend on the quality of communication between the individuals.

In birds, vocalizations exchanges lay at the heart of pair bond formation and courtship (Marler & Slabbekoorn, 2004; Tobias, Gamarra-Toledo, Garcia-Olaechea, Pulgarin, & Seddon, 2011), but vocal interactions may also function in partner's recognition (Beer, 1971; Marzluff, 1988; Robertson, 1996; Vignal, Mathevon, & Mottin, 2008), pair bond maintenance (Beletsky & Orians, 1985), foraging behaviour (Evans & Marler, 1994; Gyger & Marler, 1988), vigilance against predators (Colombelli-Negrel, Robertson, & Kleindorfer, 2011; Krechmar, 2003; Yasukawa, 1989; McDonald & Greenberg, 1991; Tobias & Seddon, 2009), and incubation of eggs and nestling provisioning (Gorissen & Eens, 2005). Some species even exhibit highly synchronized vocal duets between mates (Benedict, 2008; Dahlin & Benedict, 2013; Farabaugh, 1982; Hall, 2004, 2009).

Mates can use acoustic communication while in visual contact or when the visual contact is disrupted. Thus, there is a possibility of the amount of visual contact affecting acoustic communication during contact maintenance. Some habitat characteristics or long distances between individuals may limit the efficacy of visual communication and therefore favour contact maintenance via acoustic cues. Female Steere's liocichlas (*Liocichla steerii*) are more likely to answer their mate's song and to engage in song duets in dense forest habitat compared to open agricultural habitat (Mays, Yao, &Yuan, 2006). In the black-bellied wren (*Pheugopedius*

fasciatoventris), birds answer their mate's song more often when the mate is close, and song answering facilitates approach and direct contact (Logue, 2007). In the common marmoset (*Callithrix jacchus)*, visually occluded individuals engage in a reciprocal exchange of long-distance contact calls, a sequence called antiphonal calling (Miller & Wang, 2006), and the acoustic structure of the contact calls depends on the possibility of visual contact (Schrader & Todt, 1993). Thus, when visual contact is lost, acoustic communication seems to compensate at least part of that loss and to become more accurate: partners' respond to each other more systematically, more regularly and with specific acoustic signals. When visual contact is lost partners may be more motivated to find each other. Therefore, even if the predation risk is increased they may be more active in the acoustic channel because it becomes the central channel of communication. They also may concentrate more to hear each other in order to find each other, or they may be more efficient because they only have one channel to focus on.

In order to study the impact of the loss of visual contact on acoustic communication, we used the zebra finch (*Taeniopygia guttata*), a well studied monogamous passerine that forms lifelong pair bonds (Zann, 1996). In the wild, partners are inseparable even outside of the breeding season (McCowan, Mariette, & Griffith, in press), except during situations like incubation where only a single bird can effectively incubate alone. Even during incubation they maintain a close relationship and will act as sentinel for each other whilst carrying out the relatively vulnerable task of sitting alone in the nest (Elie et al., 2010; Mainwaring & Griffith 2013). When separated, zebra finch pairs show increased stress hormone levels as well as alterations in their behaviour that are reversed by reunion with the partner, responses considered characteristic of social bonding (Remage-Healey, Adkins-Regan, & Romero, 2003). Established pairs are able to respond quicker to an opportunity to breed (Adkins-Regan & Tomaszycki 2007), and during chick rearing, nest visits are synchronized between partners, with highly synchronized pairs achieving greater reproductive success (Mariette & Griffith, 2012, 2015). In domestic birds, foster chicks raised by parents with similar personality

traits show higher body mass and condition (Schuett, Dall, & Royle, 2011), suggesting that behavioural matching between partners could enhance parental care. The zebra finch is thus a good model species to study pair coordination and synchronization and how it potentially improves with pair bond duration. In addition, zebra finches use a large repertoire of calls during social interactions (Zann, 1996). Male and female can recognize their mates using calls only (Elie et al., 2010; Vignal, Mathevon, & Mottin, 2004; Vignal et al., 2008) and partners emit coordinated vocal duets at the nest during breeding that may help in maintaining the pair bond and coordinate brood care (Elie et al., 2010). During foraging, mates keep constant acoustic contact even when visually separated (Zann, 1996). Zebra finch mates thus remain highly coordinated in several situations where calls are involved.

Our main prediction is that partners lacking visual contact will depend more on the acoustic channel and show a better coordination in their vocal interactions. We tested this hypothesis using a protocol of separation and reunion with graded opportunity of contact which was composed of four stages: (1) partners were first separated in two acoustically isolated rooms; (2) they were allowed to be within acoustic contact at long distance and without visual contact; (3) they were reunited at close distance but still without visual contact; (4) partners were allowed to hear and see each other at close distance. The vocal activity of each bird was recorded throughout the protocol. We also recorded birds in a baseline condition, *i.e.* birds being at close distance with both visual and acoustic contact, that allowed us to characterise 'classical passive' calling behaviour, *i.e.* without perturbation. Using automatic detection/extraction algorithms, we obtained the detailed calling activity and the temporal dynamics for each individual in each condition.

We studied three sets of measures to describe the calling behaviours in different conditions. First, we focused on the call rate and time spent calling which merely depict for each bird a global and general vocal activity. Next we performed an analysis of the dynamic of calling activity in which the temporal synchrony (or lack of it) in calling activity between mates was studied by computing

the temporal cross-correlation between male and female calling signals. Then, in order to study the turn-taking sequences of the two partners with and without visual contact we chose to use Markov chains. This is a model in which the probability of being in one state (here who is emitting a call) depends only on the probability of the previous state (who emitted the last call). This model has been previously used to characterize sequences of songs syllables in birds (Kershenbaum et al., 2014), as well as human conversations: in face-to-face situation or on the phone when visual contact is not possible (Ten Bosch, Oostdijk, & de Ruiter, 2004; Wilson, M. & Wilson, T. P., 2005). Here this model is used for the first time to study the acoustic communication between partners from a new viewpoint, *i.e.* by exploring the dynamics of their acoustic exchanges.

The last two sets of measures – temporal cross-correlation and Markov chains dynamics – can together characterize important components of a pairs acoustic-dominated communication and we expect them to be refined when visual cues are absent. Finally we studied the impact of mates' history - previous breeding experience and pairs' origin (wild type or domestic) – on the vocal interaction dynamics of different pairs.

MATERIALS & METHODS

Experimental Procedure

Subjects and housing conditions

The birds used for this study were zebra finches (*Taeniopygia guttata*).

One first group of birds (25 pairs) was used for the separation/reunion protocol. In this first group, half of the animals were domestic birds bred in our colony (12 pairs), the other half were wild-type birds (13 pairs). The domestic birds had been bred in our facility for at least three generations (Tschirren et al. 2009). The captive wild type birds were either taken under licence from Sturt National Park (northwest New South Wales, Australia) in September 2007 using mist nets (Pariser,

148 Mariette, & Griffith, 2010), or were direct descendants of these wild birds, and either first or
149 second-generation captive bred.
150 Domestic and wild type birds were housed separately in two outdoor aviaries (10 X 8 m and 2 m
151 high), each containing between 30 and 50 birds. Each aviary was provided with ad libitum
152 commercial finch seeds, water, cuttlefish bones, grit, sprouted seed, two heat lamps and nestboxes.
153 We selected 20 pairs by directly observing the aviaries for three consecutive days during two hours
154 so as to detect pairs using four specific behaviours (Zann, 1996): nestling rearing (birds raising
155 chicks together), clumping (birds perching side by side in contact), allopreening (one bird preening
156 the feathers of the other one), nest sharing (birds sharing the same nest). Breeding activity in the
157 two outdoor aviaries had been monitored for a year prior to the experiment. This allowed us to
158 determine the previous reproductive success of the pair. The five remaining pairs (3 wild type and 2
159 domestic) were formed by randomly putting together a male and a female in the same individual
160 cage for one month prior to the experiment. One week before the beginning of the experiment, the
161 effectiveness of these five pair bonds was verified by observing clumping and allopreening.
162 All pairs were caught in the aviaries and then housed in individual cages (1 pair per cage, height =
163 40 cm, width = 75 cm, depth = 47 cm) stored in the same rearing room for the duration of the
164 experiment.
165
166 Another group of birds, naive to the experimental protocol, was used for the protocol in baseline
167 condition. This group was part of the colony of European Domestic zebra finches bred at the ENES
168 laboratory, University of Saint-Etienne and comprised 11 pairs. Finch seed, cuttlefish bone and
169 water were provided ad libitum and salad once a week. The temperature was maintained around 23–
170 25 °C and the photoperiod was 14L/10D. Pairs were bred in private cages (dimensions
171 40x40x25cm) put in a same room.

172 Pairs from both groups of birds had been formed at least one month before the experiment and we
173 checked if partners actually behaved as a pair using regular proxies (clumping/allopreening) used to
174 identify pairs in this species (Zann 1996).
175
176
177 *Separation – reunion protocol*
178 The day before the experiment, each pair was moved from the rearing room to the experimental
179 room and placed in a separation cage. The morning of the experiment, two webcams (logitech HD
180 pro C910) and two microphones (AKG C 417 Clip-on Microphones, one per half cage) connected
181 to a recorder (zoom H4n) were activated to monitor the birds' locomotor and vocal behaviours
182 during the whole experiment. Two sessions were run in the same morning in two different
183 experimental rooms, allowing us to record two pairs per day. Each day, the first session began at
184 8:00 am and the second session at 10:00 am. Each session lasted one hour. Wild type and domestic
185 pairs were randomly chosen to be recorded during the first or the second session. Partners were
186 physically separated using two partitions placed in the middle of the experimental cage, which
187 allowed separating the cage into two separate sections. The newly independent sections were then
188 moved in two other independent rooms separated by 6 m and two heavily insulated doors: each
189 partner was then placed in a new room, visually and acoustically isolated from its mate (Fig.1).
190 After 30 minutes of separation (Isolated), the doors from the independent rooms were opened for 10
191 minutes, allowing acoustic contact at long distance between the birds but preventing visual contact
192 (Far No Visual). This situation was suitable for the exchange of distance calls between partners.
193 Each bird was then removed from its room and placed back in the first experimental room so that
194 partners were both moved and placed in the same room again, at close distance but without visual
195 contact, for 10 minutes (Close No Visual). Finally, the partitions were removed and both acoustic
196 and visual contact were permitted during 10 minutes (Reunion).
197

198 *Determination of pair history*

199 We had two pieces of information about the pairs' history: the origin of the pair as wild-type (Wild)
200 or from domestic stock (Dom), and the previous breeding experience of each pair (as the breeding
201 experience of wild-type birds with their potential previous partner in the wild was unknown),
202 stating whether pairs successfully reared offspring (Offspring) or not (No Offspring)

203

204 *Pair recordings in baseline condition*

205 In this second protocol we recorded pairs in a baseline condition. The day before the experiment,
206 each pair was moved from the rearing room to the experimental room and each bird was placed in a
207 cage, with one microphone per cage. Microphones (Audio Technica AT8531) were connected to a
208 recorder (zoom H4n). We recorded vocal exchanges during a long period (6 hours) (9-10 am to 15-
209 16 pm) to have the opportunity to study vocal dynamic in a baseline condition.

210

211 *Call extractions*

212 Vocalizations were extracted from recordings using in-house software. These programs were
213 written in python (www.python.org) by authors H.A.S. and M.S.A.F using open-source libraries.
214 This software accuracy was tested, confirmed and used in a previous study (Elie, Soula, Mathevon,
215 & Vignal, 2011). All methods are described in this previous study and we summarize them here.
216 Vocalization detection is a pipeline of three stages. The first process was a simple threshold-based
217 sound detection based on a high-pass filtered energy envelope (1024 samples FFT; 441 Hz
218 sampling; cut-off frequency: 500Hz). During the second stage, each sound whose peak was
219 extracted was reconstructed by exploring the two sides and keeping area with energy higher than
220 10% of the peak. Thus, each event was either lengthened or shortened to obtain the same amplitude
221 range during the event. This allows a good estimate of the call duration. The third stage simply
222 merged overlapping waveform segments. Together, the three stages produced start, end, and
223 duration values for each sound event detected in the recording.

Two additional stages were added for this study in order to assign each call to its emitter and also remove cage or wing noises. The first additional stage removed double calls, *i.e.* calls emitted by one bird and recorded by its microphone but also recorded by the microphone of the other bird of the pair (only in Far No Visual, Close No Visual and Reunion conditions) by using energy and delay differences. The second stage removed cage or wings noises using a machine learning process. We trained a supervised classifier using a data set composed of 4500 random extracted sounds from all of our data. Each sound was classified by one expert (MSAF) as "call" or "non-call". The classification was performed on the spectrogram of the sounds sliced in equal parts using 55 parameters. More precisely, the spectrogram matrix was first reduced to the frequencies of interest – between 500Hz and 8kHz. To obtain the same size for all calls that have different durations, we sliced the temporal axis into 5 parts. We sliced the frequency range into 11 parts. The average value was taken to compute each entry of the reduced matrix (of size 11 by 5). This matrix will be seen as a vector of 55 parameters. We trained a Random Forest classifier (Breiman, 2001) with 1500 sounds. This classifier had an overall rate of error below 10% of the remaining 3000 sounds. We then applied an important manual verification to the extracted call sequences.

This procedure allowed us to extract two types of calls from the zebra finch repertoire: tet calls *i.e.* soft and short harmonic stacks with almost no frequency modulation (Zann, 1975; 1993), and distance calls *i.e.* complex sound consisting of a harmonic series modulated in frequency as well as amplitude (Zann, 1996). Because we were interested in the dynamic of the exchange only, we decided to pool the two types of calls in the following analyses.

Ethical note

The first group of birds was bred at the ENES laboratory, University of Saint-Etienne with the Autorisation du ministère français de la recherche, licence number 42-218-0901-38 SV 09. The second group was bred at the Macquarie University with the Animal Research Authority reference number 2010/053-5.

250 Data Analysis

251 We separated the analysis into the three parts described below: vocal activity, cross-correlation and

252 Markov analysis.

253 *Vocal activity*

254 We calculated general parameters such as call rate (number of calls per minute), cumulative number

255 of calls (total number of calls emitted from the beginning of the experiment at a given time), and

256 time spent calling (duration between the first and the last vocalization as a percentage of total

257 recording time). We also looked at the correlation between male and female call rates (Fig. 2(c)):

258 for each pair we have the male call rate on the x-axis and the female call rate on the y-axis, and

259 therefore each point represents a pair.

260

261 *Cross-correlation: Temporal analysis of male-female calling activity*

262 We computed the cross-correlation between male and female calling signals. A calling signal is a

263 temporal description of the calling emission and is defined as a function of time *t* that is 1 if the bird

264 is emitting a sound at *t* and zero otherwise. The sampling frequency was set to 200Hz (5ms bins).

265 For example if, for one part of a calling signal of 75ms, we obtain 0 0 0 1 1 1 1 1 1 1 1 1 0 0, it

266 means that during the first 15 ms (3*5ms) the bird is silent, then this bird emits a call of 50ms

267 (10*5ms) length, before it goes back to silence for 10 ms. S_{male} stands for the male signal and S_{female}

268 for the female signal. We computed the cross-correlation (cc) of these two signals (S_{male} and S_{female})

269 with the following formula:

270 $cc(T) = < (S_{male}(t) - mean(S_{male}))(S_{female}(t+T) - mean(S_{female}))>$.

271 With the normalization step, we have: $CC(T) = cc(T)/cc(0)$

272 where CC is the normalized cross-correlation, T the time delay, and S_{male} and S_{female} the male and

273 female signals as functions of t (time).

To compare cross-correlation between conditions, we computed the extreme of CC as a function of the delay T. One maximum (peak) at positive time values gives information about the delay of the male's answer to the female's call and conversely for a maximum at negative time values.

We measured several parameters on the normalized cross-correlation functions: maximum peak height, each peak height and time (for both negative and positive time delays), the area under curve, as well as the duration with curve above 0.1. The area under curve is an indicator of the variability in answer delays. The duration with curve above 0.1 is the total time interval where the cross-correlation is higher than 0.1 and represents the temporal correlation duration of vocal exchanges.

Markov analysis: Sequential analysis of male-female calling activity

As we found a strong correlation in partners' vocal activities, we expected that the vocal dynamic within pairs would present a long-term memory. To test this hypothesis we used Markov chains, a model in which the probability of being in one state (here emitting a call) depends only on the probability of the previous state (who emitted the last call), *i.e.* a model with a very short term memory. Consequently, if the vocal dynamic present a long-term memory the Markov model would be a poor predictor of the data.

Calling sequences were simply transformed into an array of M (Male call) and F (Female call) indicating the emitter's identity (e.g. MMFMFMFF). Assuming two states M and F, the call sequence can be viewed as a stochastic process that "jumps" from state to state. With the Markov hypothesis the emitter identity depends only on the previous emitter according to a transition probability (for example the probability of having an M (male call) after a F (female call)). More precisely, a Markov matrix of size 2*2 depicts the probability of jumping from one identity to the other: in this matrix, an entry at line *i* and column *j* is the probability when emitter is *i* (M or F) that the next emitter will be *j* (M or F). By construction, this matrix reproduces both the average number of calls for each individual but also the first order transition. Sequences that induce cyclic pattern like MMMFMMMFMMMF - always three M followed by F - are not Markovian because the

sequence memory is longer than one step (here it is four steps). On the other hand, sequences with a perfect alternation (MFMFMFMF) are Markovian because the probability of having an *i* depends only on the previous state: 1 if the previous state was *j* and 0 if the previous state was *i*. Totally random sequences of M and F would be Markovian, because the probability of having an *i* after a *j* (equal to 0.5) does not depend on previous states. In the latter case, by chance, we could obtain long series of M (or F) but the likelihood of such sequences occurring randomly will decrease exponentially (with the length of the series). Therefore, the statistics of series of M and F would follow a particular structure if the sequence was Markovian (here the statistic is the autocorrelation, see below).

To assess whether or not, the calling sequences are akin to a Markov model, we produced artificial call sequences based on characteristics given by the real sequence Markov matrix. To take the variability of sequences into account in this comparison, these artificial sequences are the same length as our real sequence. Therefore we can compare the artificial sequences statistics (see below) with the real sequence counterpart: we computed the real sequence autocorrelation over a signal consisting of 0 (presence of male call) and 1 (presence of female call).We then compared it to the theoretical autocorrelation of a Markov chain analytically computed as λ^T were λ is the second eigenvalue of the Markov matrix (the first eigenvalue is 1) and T is the time delay. For each time step, we located the real data's autocorrelation value in the empirical distribution of all autocorrelation values from the artificial sequences. We tested if our autocorrelation value (from the real sequence) was likely to belong to this distribution. For that we used the cumulative distribution and obtained the p-value corresponding to our real autocorrelation value. If the p-value was higher than α (5%) then there was no reason for rejecting the Markov model as a good approximation of this sequence.

Statistics

All statistical tests were performed using R software (R Core Team, 2014) and python (www.python.org).

Vocal activity

For the general parameters of vocal activity (call rate and time spent calling), as the distributions did not allow to group all factors in a single model, we used independent tests for each factor. A fit to the Normal distribution was tested using the Shapiro test. When comparing two groups, if normality was confirmed, homoscedasticity was tested using the Fisher, and if not, the Fisher test with permutation from the "RVAideMemoire" package was used (allowing non-normal data). When comparing more than two groups, the Bartlett test was used when normality was confirmed (we always had at least 4 individuals per group), or Bartlett test with permutations in the other case (non-parametric, "RVAideMemoire" package).

We had four paired conditions (Isolated, Close No Visual, Far No Visual and Reunion with the first zebra finch group), and one unpaired condition compared to the others (Baseline, with the second zebra finch group). First, when comparing between paired conditions we use either Student test for paired data (if only two means compared) or ANOVA for repeated measures (if homoscedascicity) or Friedman test (if heteroscedasticity) to test global differences between all four conditions. Wilcoxon pairwise signed rank test was then used for posthoc pairwise comparisons.

Then, when comparing between two unpaired groups, if normality was confirmed we used either Student test (if homoscedasticity) or Student test with Welch correction (if heteroscedasticity). If normality was not confirmed, Student test with permutations (if homoscedasticity) or Man-Whitney-Wilcoxon test (if heteroscedasticity) were used.

Linear model selection

From the most complex model (interactive model), simplifications were performed. When comparing models we chose to use the AICc (second order information criterion) that takes into account sample size by increasing the relative penalty for model complexity with small data sets.

The AICc converge to AIC when sample sizes increase. For each parameter, statistics resulting from the best model are presented, i.e the model with the lower AICc. P-values were computed using multiple comparisons between conditions with Tukey contrast ('glht' function of 'multcomp' R package).

Correlation between male and female general activity

We performed linear models including all factors (condition = 5 levels: Isolated, Far No Visual, Close No Visual, Reunion, Baseline, Offspring = 2 levels: Offspring (previous breeding experience), No Offspring (no previous experience) and Type = 2 levels: Wild (wild type) and Dom (domestic)) and the pair identity as random factor. We selected the following linear models: female call rate ~male call rate * condition + 1|pair and female time spent calling ~male time spent calling * condition + 1|pair. As the interactions between the two factors were significant, we studied the influence of male call rate and time spent calling in each condition separately. Detailed results of the models are shown in Table A1.

Probability of emitting at least one call

The probability of emitting at least a call was studied using a generalized mixed model with a binomial family, with a 0/1 response (0 if the bird did not emit one call during the recording, 1 if the bird emitted at least one call). The following model was selected: probaOneCall ~ condition * Offspring + 1|pair. As the interaction between the two factors was significant, we studied the influence of Offspring on the probability of emitting a call in each condition separately. Detailed results of the model are shown in Table A2.

Markov analysis

For each time step we had the success or failure for our real data to belong to the theoretical distribution of Markov (0 if data did not belong to the distribution, 1 if data belong to the distribution). Birds from the baseline group had all the same previous breeding experience (Offspring) and they were all domestic, so we could not include them in a global model with Offspring and TYPE factors. As a consequence we first built a model including the condition as a factor (markov fit ~ condition + 1|pair, generalized linear mixed models with binomial family). This model was validated and it was thus possible to interpret the results. However we also wanted to know if the previous breeding experience (Offspring/No Offspring) and the type (Wild/Dom) had an influence on the Markov fit. We built generalized linear mixed models with binomial family including all factors (condition, Offspring, TYPE) and selected the following: markov fit ~ condition * TYPE + 1|pair. As the interaction between factors was significant, we studied the influence of TYPE in each condition. All binomial models were checked as explained in Model validation section. Detailed results of the models are shown in Table A3.

Model validation

Before being interpreted each model was checked, paying particular attention to their residuals. For binomial models, we used five relevant plots from custom-written codes (Collett, D., 1991; Atkinson, A.C., 1981) to test the validity. First, with the graph of standardized deviance residuals we checked the residuals mean homogeneity, and with the graph of absolute value of standardized deviance residuals we checked the residuals variance homogeneity. For both plots we only checked if the residuals were between -2 and 2: because of the binary nature of date (and contrary to classical linear models), non-homogeneously distributed residuals are no necessarily reflecting an inappropriate model. The model hat matrix was then extracted and its diagonal coefficients (h_i) enabled us to check the general influence of observations on the model fit to data. The threshold for h_i values is 2*mean(h_i). The Cook's distance gave us information about the influence of each observation on the parameter estimation, and had to be lower than 4/n with n the number of

observations. Finally, we build the half-normal plot (Atkinson, A.C., 1981) *i.e*. standardized deviance residuals as a function of the half-normal distribution quantiles with simulated envelope. If data points were included in the envelope, the linear predictor was correct.

RESULTS

Vocal activity: Call rate increases with decreasing distance and with visual contact opportunity.

The five conditions triggered different vocalization behaviours of the pairs as they significantly affected both call rate and time spent calling (Friedman test: call rate: $\chi^2_3 = 37.75$, $P < 0.001$; time spent calling $\chi^2_3 = 43.41$, $P < 0.001$). Call rates in the Close No Visual and Reunion conditions were higher than in all other conditions (Fig. 2(a)). Time spent calling was higher in the Close No Visual condition than in the Isolated and Far No Visual conditions (Fig. 2(b)), and was even higher when visual contact was possible (Reunion and Baseline conditions) (see Table 1).

Birds in Isolated and Far No Visual conditions displayed low levels of vocal activity (Fig. 2(a) and 2B). Some pairs did not emit calls at all (1 of 25 in Isolated and 10 of 25 in Far No Visual), but most pairs did, and for pairs that emitted at least one call, even though calls were few, they were spread over a large percentage of the recording time (62 ± 36 % for Isolated and 63 ± 35 % in Far No Visual). Compared to the Close No Visual and Reunion conditions, visual contact in Baseline condition was associated with significantly lower call rates (Fig. 2(a)), but high percentage of time spent calling (Fig. 2(b)). In all conditions, there was no difference between the sexes either in call rate or in time spent calling (Table 2).

Vocal activity: the correlation between male and female vocalizations increases with proximity but vanishes with visual contact.

We computed the correlation coefficient (R^2) between the call rates and time spent calling of the male and the female of each pair in each condition (linear models with significant interactions, male

call rate * condition: $F_{4,66}$=5.47, P = 0.001 and male time spent calling * condition: $F_{4,66}$=4.58, P = 0.002) (Fig. 2(c) and Fig. A1). Correlation increased significantly between the isolation condition (Table 3, Isolated, R^2 call rate = -0.02 [-0.41, 0.37], R^2 time spent calling = 0.35 [-0.052,0.65]) and conditions allowing acoustic contact only (Table 3, Far No Visual R^2 call rate = 0.76 [0.53, 0.89], R^2 time spent calling = 0.89 [0.77,0.95], and Close No Visual R^2 call rate = 0.89 [0.77, 0.95], R^2 time spent calling = 0.99 [0.99,1.0]). Thus, without visual contact, the closer male and female are, the higher the correlation of their vocal emissions. This suggests that an increase in acoustic contact probability after a separation leads to a more correlated vocal activity. This correlation was lower when visual contact was allowed (Table3, Reunion R^2 call rate = 0.21 [-0.20, 0.56], R^2 time spent calling = 0.83 [0.66,0.92], and Baseline R^2 call rate = 0.46 [-0.19, 0.83], R^2 time spent calling = 0.21 [-0.45,0.71]).

Thus, pairs showed correlated vocal activities during vocal exchanges without visual contact. This correlation decreased with visual contact, and could sign a return to a baseline condition.

Cross-correlation: Pairs' temporal coordination is different in acoustic or visual contact.

The cross-correlation of mates' calling signals significantly differed between the five experimental conditions (Fig. 3). The vocal coordination was lower when visual contact was allowed (Reunion and Baseline conditions). Without visual contact, the cross-correlation was higher but varied according to the distance between mates. In the Far No Visual context, two peaks on the cross-correlation plot (Fig. 3) indicate that male and female answered each other alternatively: the left peak (with a negative time value) reflects the fact that on average the female answered the male with a 600ms delay, and the right peak (with a positive time value) tells us that on average the male answered the female with a 350ms delay. In the Close No Visual context, one unique peak indicates that one partner (here the female) answered the other with a 40ms delay. Illustrations of these results are shown in figure A2, with an example of call emission dynamic for one pair.

453 Among all pairs, the number of peaks of the cross-correlation differed between the Far No Visual
454 and Close No Visual contexts (exact binomial test: $P = 0.007$), with two peaks being more likely in
455 the Far No Visual condition and one peak in the Close No Visual condition.
456
457 *Markov Analysis: Mates call sequence statistics are Markovian only without visual contact*
458 Here most pairs' vocal exchanges followed very closely a Markovian pattern when visual contact
459 was not allowed (Fig. 4). The fit to Markov of the call sequence was lower when visual contact was
460 possible as the Reunion and Baseline conditions each differed from both the Close No Visual and
461 the Far No Visual conditions. This difference in the Markov fit was significant between conditions
462 (generalized linear model with binomial family, Table 4). In other words, without visual contact the
463 decision to emit a call is taken on a very short memory: the identity of an emitter is well predicted
464 by the knowledge of the identity of the previous emitter only.
465
466 *Mates' history modifies vocal coordination*
467 Our results suggest that pairs correlate their vocal exchange and we wanted to assess if this capacity
468 was related to the history of the pair for some of the three measures described above – Vocal
469 activity, Cross-correlation and Markov analysis. For the first group of birds (corresponding to the
470 Isolated, Close No Visual, Far No Visual and Reunion conditions) we had information about the
471 prior breeding experience of pairs (Offspring/No Offspring indicating whether the partners had a
472 breeding experience together or not) and their type (Dom for domestic or Wild for wild-type).
473
474 Vocal activity: In the Far No Visual condition, pairs with successful breeding experience
475 (Offspring) were more likely to emit calls than other pairs (No Offspring) (Fig 5(a), generalized
476 linear model with significant interaction between condition and Offspring factors, $z = -3.21$, N=50,
477 $P = 0.001$). So, when separated and able to hear each other, pairs with prior breeding experience
478 were more likely to start a vocal exchange.

Cross-correlation: Besides, among pairs starting a vocal exchange in the Far No Visual and Reunion contexts, pairs with breeding experience presented more regularity in their delay of response to each other than unexperienced ones (cross-correlation maximum peak height for No Offspring vs Offspring, student test with Welch correction: $t_{12.77} = 2.23$, $N = 25$ pairs, $P = 0.004$ in Far No Visual and $P = 0.044$ in Reunion). In the Reunion context, temporal correlation of vocal exchanges was longer for experienced pairs (duration with a cross-correlation higher than 0.1, Wilcoxon exact rank sum test: $W = 104$, $N = 25$ pairs, $P = 0.054$).

In the Far No Visual context, temporal correlation of vocal exchanges was longer in wild-type pairs than in domestic pairs (duration with cross-correlation value higher than 0.1: student test with permutations: $t = 2.2$, $N = 13$ pairs, $P = 0.049$).

Markov fit: The difference in the Markov fit was significant between conditions, but this difference was not the same between the wild and domestic groups. Figure 5(b) shows that wild-type were more likely to fit better a Markov dynamic than domestic (Far No Visual: $z = 2.77$, $N = 12$, $P = 0.005$; Close No Visual: $z = 3.73$, $N = 23$, $P < 0.001$; Reunion: $z = 2.92$, $N = 25$, $P = 0.003$). Wild-type pairs show 78.5 ± 33.3 % points in time fitting Markov whereas domestic pairs show 58.2 ± 33.7 % points fitting Markov.

DISCUSSION

Our experiment revealed a strong correlation between mates' vocal activities that was stronger when birds were in acoustic but not visual contact. The temporal cross-correlation between a pair's vocalisations was higher when only acoustic contact was allowed (in both reunion and baseline conditions). We also found using a Markov analysis that the turn-taking sequence of the two partners was more predictable when pairs were unable to see one another. Both the temporal correlation and Markov property reveal that without visual clues the decision to emit a call is taken

505 on a very short-term basis and solely on acoustic information, indicating that birds adapt their
506 calling dynamics to cope with limited visual cues.
507
508 Taken together, these results show that partners possibly compensate the lack of visual contact by
509 improving vocal interaction to maintain an equivalent level of contact maintenance. This could also
510 indicate a slow return to the baseline situation, due to the close and assured presence of the partner.
511 This high correlation of vocal emissions in acoustic contact only between partners probably
512 indicates that the probability of response of one individual is strongly dependent on the actual
513 emission of the other. Assuming this is the case, we proposed a sequential analysis and studied the
514 turn-taking sequence using Markov chains paradigm. As we showed, most turn-taking sequences
515 showed Markov-like dynamics when acoustic contact is possible.
516
517 Similar vocal activity between partners when only acoustic contact is allowed could indicate an
518 assortative mating on partners' vocal profile, silent birds and talkative birds being mated together.
519 However, calling behaviours of the male and the female are strongly different in other conditions.
520 Thus, the similar behaviours when no visual contact is possible could more likely indicate an
521 adjustment of behaviour. This similarity between mates through an adjustment of behaviour has
522 been described previously. For example in black-bellied wrens (*Pheugopedius fasciatoventris*),
523 even if females are able to sing different types of songs, they match the song-type of their mate
524 during duets (Logue, 2006). A study of vocal learning in budgerigars (*Melopsittacus undulatus*)
525 also showed that males imitate the contact calls of their newly assigned female (Hile, Plummer, &
526 Striedter, 2000).
527
528 Vocal activities of partners during acoustic contact showed a strong correlation of call rate and time
529 spent calling and high coordination (temporal cross-correlation). These data could fit some of the
530 characteristics used to define vocal duets (Farabaugh, 1982; Hall, 2004; Wickler & Seibt, 1982),

especially in long distance acoustic contact. In our study, vocal exchanges between partners when acoustic contact only is possible could thus be seen as duet-like exchanges that participate in pair bond maintenance after a separation. However it has to be noted that coordinated calling activity does not necessarily imply a duet or a conversation. In some social contexts, birds can adjust the timing of their calls only to reduce vocal costs, and the resulting vocal activity of the group is then coordinated. For example, between parental feeding visits, barn owl (*Tyto alba*) young siblings optimize communication and adjust their call timing to avoid signal interference (Dreiss, Ruppli, Faller, & Roulin, 2015).

When partners are only in acoustic contact, turn-taking sequence dynamics are not distinguishable from a Markov chain. This is an unexpected result. Indeed, Markov chains are systems that display exponentially decreasing autocorrelation due to short-term memory. Yet, we found a strong correlation in partners' vocal activity. Long time correlation can usually be explained by oscillating behaviours (one individual repeating the same pattern with a long period such as a sequence of MMMFMMMF – 3M's followed by a F) or by long-term memory (one or both individuals recall patterns of emission far back in time). In the context of a new acoustic contact after separation this does not seem to be the case and very short memory (Markov-like dynamics) seems to be the rule when no visual contact is possible. In this condition, the birds' decision to emit a call seems to depend only on the previous emitted call, and this indicates the presence of a discussion rather than a proper rhythm of calling emission for each partner.

There is some relationship to human discussion behaviours. In social science, conversation experts suggest that humans agree with implicit conversational rules that determine the optimal moment to alternate the speaker (Duncan, 1972; Sacks, Schegloff, & Jefferson, 1974). Turn-taking involves highly coordinated timing, with short response times and dynamics depending only on the last emission (Choudhury & Basu, 2004; Takahashi, Narayanan, & Ghazanfar, 2013; Wilson, M. & Wilson, T. P., 2005). Ten Bosch et al. (2004) studied the differences in turn-taking behaviour between face-to-face and phone conversations in humans. Phone conversations (thus relying only

on acoustic cues) show shorter pauses than face-to-face dialogues, and furthermore the variability in pauses duration is larger in face-to-face dialogues. This was also the case in our study of zebra finches: during acoustic contact, delays of response between partners are extremely precise whereas in visual contact the answer delay is much more variable. This confirms that for birds such as for humans, the context of conversation seems to be an important factor for the timing aspects of turn-taking. A possible explanation for this phenomenon is that in face-to-face conversations individuals have several ways to convey information and to let know to their partner they are still involved in the conversation, without having to use acoustic signals. During the conditions allowing visual contact (reunion and baseline) it would be interesting to focus on visual signals between partners in addition to acoustic ones. We found that a turn-taking sequence made of acoustic signals only is not Markovian in this case, but a sequence made of both visual and acoustic signals could reveal that birds' decision to emit a signal (visual or acoustic) depends only on the previous emitter, *i.e.* that this new type of sequence is Markovian.

Our results suggest that pairs very well coordinate their vocal exchange and we wanted to assess if this capacity was related to the history of the pairs. In our experiment some individuals remained silent during long distance acoustic accessibility and this can be explained by the protocol: whenever one individual emitted a call the other one could hear it and start a vocal exchange. On the other hand, there was no reason to detect that the other was within earshot without trying and emitting a call. We show that the pairs' history did indeed play a role when such a contact opportunity emerges: experienced pairs were more likely to start calling to elicit answer from partner than inexperienced ones.

This may be explained by several causes. It is possible that experienced pairs show more reliable vocal recognition between partners. A study in mandarin voles (*Microtus mandarinus*) and root voles (*Microtus oeconomus*) suggests that the intensity of mate recognition by sniffing or licking is related to familiarity degree (Zhao, Y., Tai, Wang, Zhao, X., & Li, 2002). Comparing discrimination of mate versus non-mate calls of birds from experienced pairs and inexperienced

584 pairs would test this hypothesis. In this study, the acoustic contact between pairs occurred after a
585 separation, thus it represents a vocal reunion in a non-reproductive context and could contribute to
586 pair bond maintenance. Inexperienced pairs are quieter in this context perhaps because their bond is
587 weaker and partners are less motivated to maintain contact. It is possible that previous breeding
588 experience allowed partners to exhibit many vocal exchanges, especially during duets at the nest
589 (Elie et al., 2010) and this could have led to a talkative and coordinated pair. As mate separation
590 results in an increase in corticosterone concentrations in zebra finches (Remage-Healey et al.,
591 2003), it is possible that the stress of isolation and visual separation differs between experienced
592 and inexperienced pairs and provokes different levels of vocal activity. This remains to be
593 investigated.
594
595 We showed that vocal interactions of wild-type birds fit better with a Markov model than domestic
596 ones when in acoustic contact. It has been shown that even if there is no evidence for a bottleneck
597 due to domestication of zebra finches, captive populations have lost some of the genetic variability
598 present in the wild (Forstmeier, Segelbacher, & Mueller, 2007). Our two groups of birds thus have
599 genetic differences that could explain their different call dynamics. Parameters used in previous
600 studies to compare wild and domestic zebra finch behaviour have not revealed any significant
601 difference between these populations (Tschiren et al., 2009). Besides, to our knowledge, no element
602 could explain vocal dynamic differences between wild and domestic zebra finches. However,
603 Honda and Okanoya (1999) showed that White-backed Munia (*Lonchura striata*) and its domestic
604 strain, the Bengalese Finch (*Lonchura striata var. domestica*) present differences in both the
605 acoustical properties and the temporal aspect of their songs. This could also be the case in zebra
606 finch call dynamics. Additional experiments are needed on this point.
607
608 Here we confirm that classical metrics like mean or coefficient of variation are not always sufficient
609 for the study of animal vocal interaction sequences (Kershenbaum et al., 2014). We show that short-

610 term memory model like Markov can explain vocal exchanges dynamics in a particular context (no
611 visual contact), but long-term memory dynamics should be studied in various contexts in the future.

612 Zebra finches form life-long pair bonds, with low levels of extra-pair paternity (Birkhead, Burke,
613 Zann, Hunter, & Krupa, 1990; Zann, 1996; Griffith et al. 2010), and show high coordination of pair
614 activities during and outside reproduction (Mariette & Griffith 2012; Mainwaring & Griffith 2013).
615 Thus, zebra finches show one of the strictest social and reproductive monogamy in birds. Here we
616 show that, without visual cues, a form of synchronization and coordination of the pair is expressed
617 through the strong correlation of partners' calling activities. This coordination decreases with visual
618 contact as birds' vocalizations return to an individual baseline dynamic.
619 This study provides new insight into the question of how birds can adapt their calling dynamics to
620 cope with limited visual cues. Without visual contact, pairs' vocal activity is highly correlated and
621 the decision to emit a call is taken only on acoustic information and on a very short-term basis. That
622 way, this calling dynamics may increase the amount of information by decreasing the uncertainty
623 when visual contact is not possible and when acoustic transmission becomes the only channel of
624 communication.
625
626
627 ACKNOWLEDGMENTS
628
629 This work was supported by an ANR grant to CV (French Agence Nationale de la Recherche,
630 project) 'Acoustic Partnership'), a joint NSF/ANR – CRCNS grant 'AuComSi' for MSAF and
631 HAS, an IUF grant to CV (Institut Universitaire de France) and an Australian Research Council
632 grant (DP0881019) to S.C.G. E.P. was supported by an International Macquarie University
633 Research Excellence Scholarship and a CMIRA grant awarded by the French Rhônes-Alpes region.
634 We are grateful to Colette Bouchut and Nicolas Boyer for their help at the ENES lab, and to Kate
635 Buchanan (Deakin University) for discussions.

636 REFERENCES

637

638 Adkins-Regan, E. & Tomaszycki, M. (2007). Monogamy on the fast track. *Biology Letters, 3*, 617-
639 619.

640

641 Atkinson, A. C. (1981). Two graphical displays for outlying and influential observations in
642 regression. *Biometrika*, 13-20.

643

644 Beer, C. (1971). Individual recognition of voice in the social behaviour of birds.
645 *Advances in the Study of behaviour, 3,* 27-74

646

647 Beletsky, L. D. & Orians, G. H. (1985). Nest associated vocalisations of female redwinged
648 blackbirds, *Agelaius phoeniceus*. *Zeitschrift für Tierpsychologie, 69*, 329e339.

649

650 Benedict, L. (2008). Occurrence and life history correlates of vocal duetting in North American
651 passerines. *Journal of Avian Biology, 39(1)*, 57-65.

652

653 Birkhead, T., Burke, T., Zann, R., Hunter, F., & Krupa, A. (1990). Extra-pair paternity and
654 intraspecific brood parasitism in wild zebra finches taeniopygia guttata, revealed by dna
655 fingerprinting. *Behavioral Ecology and Sociobiology, 27(5)*, 315–324.

656

657 Black, J. M. and Hulme M. (1996). Partnerships in Birds: The Study of Monogamy, Oxford
658 University Press.

659

660 Black, J. M. (2001). Fitness consequences of long-term pair bonds in barnacle geese: monogamy in
661 the extreme. *Behavioral Ecology, 12(5),* 640-645.
662
663 Breiman, L. (2001). Random Forests. *Machine Learning, 45(1)*, 5–32.
664
665 Choudhury, T., & Basu, S., "Modelling conversational dynamics as a mixed memory markov
666 process", in *Proceedings of Neural Information Processing Systems Conference (NIPS)*, December
667 2004.
668
669 Collett, D. (1991). Modelling binary data. Chapman & Hall.
670
671 Colombelli-Negrel, D., Robertson, J., & Kleindorfer, S. (2011). Risky revelations: Superb Fairy-
672 wrens Malurus cyaneus respond more strongly to their mate's alarm song. *Journal of Ornithology,*
673 *152(1),* 127-135.
674
675 Dahlin C. R. & Benedict L. (2013). Angry Birds Need Not Apply: A Perspective on the Flexible
676 form and Multifunctionality of Avian Vocal Duets. *Ethology, 119*, 1-10
677
678 Dreiss, A. N., Ruppli, C. A., Faller, C., & Roulin, A. (2015). Social rules govern vocal competition
679 in the barn owl. *Animal Behaviour, 102*, 95-107
680
681 Duncan, S. (1972). Some signals and rules for taking speaking turns in conversations. *Journal of*
682 *Personality and Social Psychology, 23,* 283–292.
683

684 Elie J. E., Soula H. A., Mathevon N., & Vignal C. (2011). Dynamics of communal vocalizations in
685 a social songbird, the zebra finch (*Taeniopygia guttata*). *Journal of the Acoustical Society of*
686 *America, 129(6),* 4037-46.
687
688 Elie, J. E., Mariette, M. M., Soula, H. A., Griffith, S. C., Mathevon, N., & Vignal, C. (2010). Vocal
689 communication at the nest between mates in wild zebra finches: a private vocal duet? *Animal*
690 *Behaviour, 80(4),* 597-605.
691
692 Ens, B. J., Safriel, U. N., & Harris, M. P. (1993). Divorce in the long-lived and monogamous
693 oystercatcher, Haematopus ostralegus: incompatibility or choosing the better option?" *Animal*
694 *Behaviour, 45(6),* 1199-1217.
695
696 Evans, C. S. & Marler, P. (1994). Food calling and audience effects in male chickens, Gallus gallus:
697 their relationships to food availability, courtship and social facilitation. *Animal Behaviour, 47(5),*
698 1159-1170.
699
700 Farabaugh, S. M. (1982). The ecological and social significance of duetting. *Acoustic*
701 *communication in birds, 2,* 85-124.
702
703 Forslund, P & Pärt, T. (1995). Age and reproduction in birds - hypotheses and tests. *Trends in*
704 *Ecolology & Evolution, 10,* 374-8.
705
706 Forstmeier, W., Segelbacher, G., & Mueller, J. C. (2007). Genetic variation and differentiation in
707 captive and wild zebra finches (*Taeniopygia guttata*). *Molecular Ecology, 16,* 4039–4050.
708

709 Gorissen, L. & Eens, M. (2005). Complex female vocal behaviour of great and blue tits inside the
710 nesting cavity. *Behaviour, 142,* 489-506.
711
712 Gyger, M. & Marler, P. (1988). Food calling in the domestic fowl, Gallus gallus: the role of
713 external referents and deception. *Animal Behaviour, 36(2),* 358-365.
714
715 Hall, M. L. (2004). A review of hypotheses for the functions of avian duetting. *Behavioural*
716 *Ecology and Sociobiology, 55(5),* 415-430.
717
718 Hall, M. L. (2009). A review of vocal duetting in birds. *Advances in the Study of behaviour, 40,* 67-
719 121.
720
721 Hile, A. G., Plummer, T. K., & Striedter, G. F. (2000). Male vocal imitation produces call
722 convergence during pair bonding in budgerigars, Melopsittacus undulatus. *Animal Behaviour,*
723 *59(6),* 1209-1218.
724
725 Honda, E. & Okanoya, K. (1999). Acoustical and Syntactical Comparisons between Songs of the
726 White-backed Munia (*Lonchura striata*) and Its Domesticated Strain, the Bengalese Finch
727 (*Lonchura striata var. domestica*). *Zoological Science, 16(2),* 319-326.
728
729 Kershenbaum, A., Bowles, A. E., Freeberg, T. M., Jin, D.Z., Lameira, A. R., & Bohn, K. (2014).
730 Animal vocal sequences: not the Markov chains we thought they were. *Proceedings of the Royal*
731 *Society B.,* 281:20141370
732
733 Krechmar, E. A. (2003). Alarm calls in duets of the white-fronted goose, *Anser albifrons*.
734 *Zoologichesky Zhurnal, 82(10),* 1239-1249.

735

736 Logue, D. M. (2006). The duet code of the female black-bellied wren. *The Condor, 108(2),* 326–
737 335.

738

739 Logue, D. M. (2007). Duetting in space: a radio-telemetry study of the black-bellied
740 wren. *Proceedings of the Royal Society B., 274,* 3005-3010
741 doi: 10.1098/rspb.2007.1005

742

743 McDonald, M. V. & Greenberg, R. (1991). Nest departure calls in female songbirds. *Condor*, *93*,
744 365e373.

745

746 McCowan, L. C., Mariette, M. M., & Griffith, S. C. (in press) The size and composition of social
747 groups in the wild zebra finch. *Emu*

748

749 Mainwaring, M. C. & Griffith, S. C. (2013). Looking after your partner: sentinel behaviour in a
750 socially monogamous bird. *PeerJ* 1:e83

751

752 Mariette, M. M. & Griffith, S. C. (2012). Nest visit synchrony is high and correlates with
753 reproductive success in the wild Zebra finch *Taeniopygia guttata*. *Journal of Avian Biology, 43(2),*
754 131-140.

755

756 Mariette, M. M. & Griffith, S. C. (2015). The adaptive significance of provisioning and foraging
757 coordination between breeding partners. *The American Naturalist, 185(2)*, 270-280.

758

759 Marler, P. R. & Slabbekoorn, H. (2004). Nature's music: the science of birdsong, Academic Press.

760

761 Marzluff, J. M. (1988). Vocal recognition of mates by breeding Pinyon Jays, *Gymnorhinus*
762 *cyanocephalus*. *Animal Behaviour, 36(1),* 296-298.
763
764 Mays, J. H. L., Yao, C.-T., & Yuan, H.-W. (2006). Antiphonal duetting in Steere's liocichla
765 (*Liocichla steerii*): male song individuality and correlation between habitat and duetting behaviour.
766 *Ecological Research, 21,* 311-314.
767
768 Miller, C. T. & Wang, X. (2006). Sensory-motor interactions modulate a primate vocal behavior:
769 antiphonal calling in common marmosets. *Journal of Comparative Physiology A, 192*, 27–38.
770
771 Pariser, E. C., Mariette, M. M. & Griffith S. C. (2010). Artificial ornaments manipulate intrinsic
772 male quality in wild-caught zebra finches (*Taeniopygia guttata*). *Behavioral Ecology, 21,* 264-269.
773
774 Remage-Healey, L., Adkins-Regan, E., & Romero, L. M. (2003). Behavioral and adrenocortical
775 responses to mate separation and reunion in the zebra finch. *Hormones and Behavior, 43,* 108–114.
776 (doi:10.1016/S0018-506X(02)00012-0)
777
778 Robertson, B. C. (1996). Vocal mate recognition in a monogamous, flock-forming bird, the
779 silvereye, *Zosterops lateralis*. *Animal Behaviour, 51(2),* 303-311.
780
781 Ryan, K. K. and Altmann J. (2001). Selection for male choice based primarily on mate
782 compatibility in the oldfield mouse, *Peromyscus polionotus rhoadsi*. *Behavioural Ecology and*
783 *Sociobiology, 50(5),* 436-440.
784
785 R Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for
786 Statistical Computing, Vienna, Austria.

787

788 Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of
789 turntaking for conversation. *Language, 50*, 696-735.

790

791 Schrader, L. & Todt, D. (1993). Contact call parameters covary with social context in common
792 marmosets, *Callithrix j. jacchus*. *Animal Behaviour, 46(5),* 1026-1028.

793

794 Schuett, W, Dall, S. R. X., & Royle, N. J. (2011). Pairs of zebra finches with similar 'personalities'
795 make better parents. *Animal Behaviour, 81(3),* 609-618

796

797 Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. a. (2013). Coupled oscillator dynamics of
798 vocal turn-taking in monkeys. *Current Biology : CB*, *23(21)*, 2162–8. doi:
799 10.1016/j.cub.2013.09.005

800

801 ten Bosch, L., Oostdijk, N. , & de Ruiter, J.P. (2004). Durational aspects of turn-taking in
802 spontaneous face-to-face and telephone dialogues. Text, Speech and Dialogue. Lecture notes in
803 Computer Science Volume 3206 pp 563-570

804

805 Tobias, J. A. & Seddon, N. (2009). Signal design and perception in Hypocnemis antbirds:
806 evidence for convergent evolution via social selection. *Evolution 63*, 3168–3189

807

808 Tobias, J. A., Gamarra-Toledo, V., Garcia-Olaechea, D., Pulgarin, P. C., & Seddon N.
809 (2011). Year-round resource defence and the evolution of male and female song in suboscine birds:
810 social armaments are mutual ornaments. *Journal of Evolutionary Biology, 24*, 2118–2138

811

812 Tschirren, B., Rutstein, A. N., Postma, E., Mariette, M., & Griffith, S. C. (2009). Short- and long-
813 term consequences of early developmental conditions: a case study on wild and domesticated zebra
814 finches. *Journal of Evolutionary Biology, 22,* 387–395.
815
816 Vignal, C., Mathevon, N., & Mottin, S. (2004). Audience drives male songbird response to partner's
817 voice. *Nature, 430(6998),* 448-451.
818
819 Vignal, C., Mathevon, N., & Mottin, S. (2008). Mate recognition by female zebra finch: Analysis of
820 individuality in male call and first investigations on female decoding process. *Behavioural*
821 *Processes, 77(2),* 191-198.
822
823 Wickler, W. and Seibt, U. (1982). Song splitting in the evolution of dueting. *Zeitschrift für*
824 *Tierpsychologie, 59(2),* 127-140.
825
826 Wilson M. & Wilson T. P. (2005) An oscillator model of the timing of turn taking. *Psychononomic*
827 *Bulletin and Review, 12(6),* 957-968.
828
829 Yasukawa, K. (1989). The costs and benefits of a vocal signal: the nest associated 'chit' of the
830 female red-winged blackbird, *Agelaius phoeniceus*. *Animal Behaviour, 38*, 866e87
831
832 Zann, R. A. (1975). Inter- and intraspecific variation in the calls of three species of grassfinches of
833 the subgenus *Poephila* (Gould) (Estrildidae). *Zeitschrift für Tierpsychologie 39*: 85-115
834
835 Zann, R. A. (1993). Structure, sequence and evolution of song elements in wild Australian Zebra
836 Finches. *Auk 110*: 702-715
837

838 Zann R. A. (1996). The Zebra Finch, a synthesis of Field and Laboratory Studies, Oxford
839 University Press.
840
841 Zhao, Y., Tai, F., Wang, T., Zhao, X., & Li, B. (2002). Effects of the familiarity on mate choice and
842 mate recognition in *Microtus mandarinus* and *M. oeconomus*. *Chinese Journal of Zoology, 48(2),*
843 167-174.

844 FIGURES LEGENDS

845

846 Figure 1: Protocol design

847 Schematic describing the separation/reunion protocol. Cages were acoustically and visually
848 separated (Isolated) then visually separated at long distance via the doors opening (Far No Visual),
849 visually separated at short distance (Close No Visual), and then visually reunited (Reunion). We
850 also recorded zebra finch pairs in a baseline condition (visual and acoustic contact at short distance)
851 during 6 hours.

852

853 Figure 2: Call rate and time spent calling analysis

854 (a) Call rate (mean number of calls per minute) per recording and for each sex. For the Isolated, Far
855 No Visual, Close No Visual and Reunion conditions, N=25 pairs, and for the Baseline condition,
856 N=11 pairs. Bars are means, and Confidence Interval 5% for non-normal data are given. Different
857 letters indicate significant differences (posthoc test after Friedman test for paired data and student
858 tests for independent data). Pairwise tests yield no significant differences either between sexes. (b)
859 Time spent calling in each condition and for each sex (duration between the first vocalization and
860 the last as a percentage of total recording time). For the Isolated, Far No Visual, Close No Visual
861 and Reunion conditions, N=25 pairs, and for the Baseline condition, N=11 pairs. Bars are means,
862 and Confidence Interval 5% for non-normal data are given. Different letters indicate significant
863 differences (posthoc test after Friedman test for paired data and student tests for independent data).
864 Pairwise tests yield no significant differences between sexes. (c) Correlation of male and female
865 call rates. Linear regression of female versus male call rates depending on the condition. Isolated R^2
866 = -0.02 [-0.41, 0.37], Far No Visual R^2 = 0.76 [0.53, 0.89], Close No Visual R^2 = 0.89 [0.77, 0.95],
867 Reunion R^2 = 0.21 [-0.20, 0.56], Baseline R^2 =0.46 [-0.19, 0.83].

868

869

Figure 3: Mean cross-correlation between male and female signals for each condition, over all pairs.

Figure 4: Percentage of fit to Markov model for male/female call sequences.

Fit to Markov model as the percentage of points in time statistically close to the theoretical Markov autocorrelation values. Bars are means, with Confidence Interval 5% for non-normal data. Different letters indicate significant differences (posthoc test after Friedman test for paired data and Wilcoxon test or student test with permutations for independent data).

Figure 5: Influence of mates history on vocal coordination.

(a) Probability to remain silent in Far No Visual condition in experienced (Offspring) vs non-experienced pars (No Offspring). Bars are means, with Confidence Interval 5% for non-normal data. (b) Percentage of fit to Markov model of wild-type (Wild) vs domestic (Dom) birds.

Fit to Markov model as the percentage of points in time statistically close to the theoretical Markov autocorrelation values. Bars are means, with Confidence Interval 5% for non-normal data.

Figure A1: Correlation of male and female time spent calling

Linear regression of female versus male time spent calling depending on the condition: Isolated R^2 = 0.35 [-0.05,0.65], Far No Visual R^2 = 0.89 [0.77,0.95], Close No Visual R^2 = 0.99 [0.99,0.1], Reunion R^2 = 0.83 [0.65,0.92], Baseline R^2 =0.21 [-0.45,0.72].

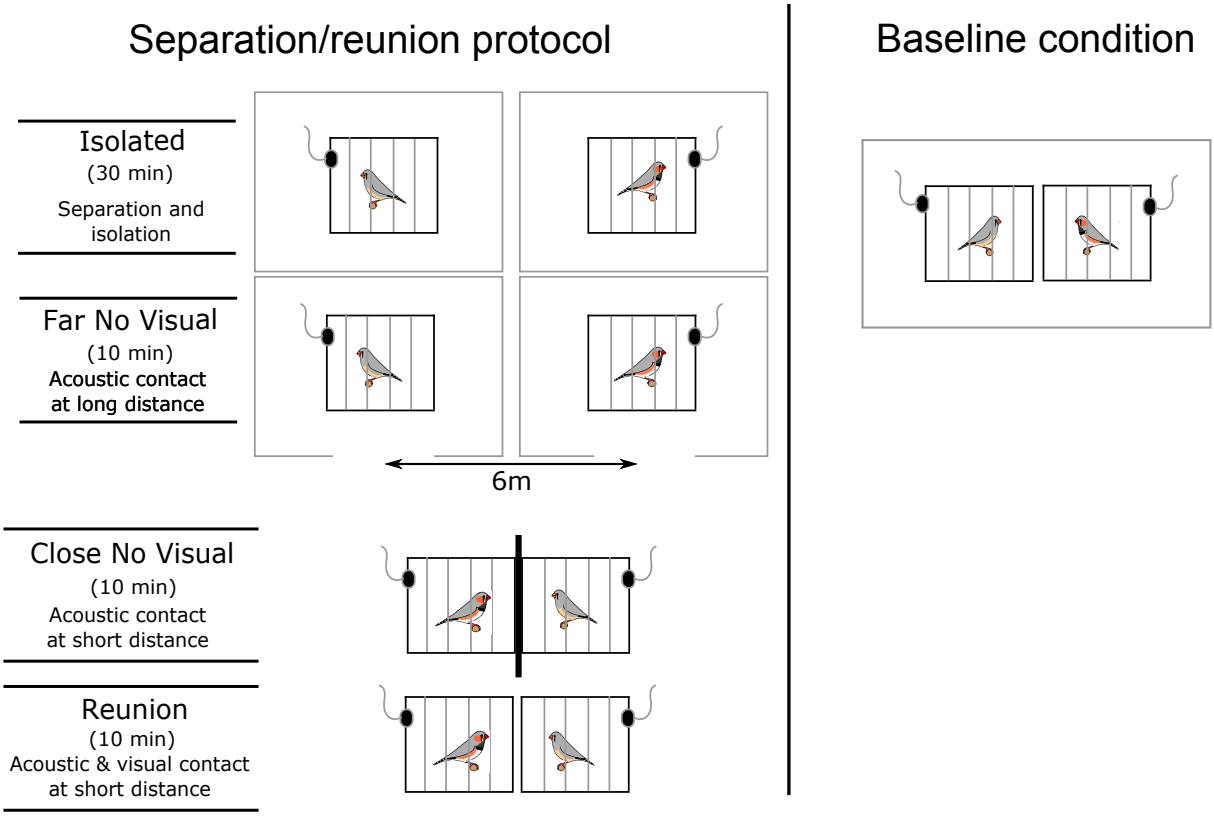
Figure A2: Example of cumulative number of calls for one male and one female and associated call timing for each condition.

(a) Cumulative number of calls. The inset shows that for the Baseline condition, we present a minute from a burst period, *i.e.* when the call rate is high. When visual contact was not possible (dotted lines), the cumulative numbers of calls for the male and the female were highly correlated, with period where both the male and the female are calling (see the arrow). When visual contact is

allowed, curves of cumulative number of calls were no longer correlated. (b) Call timing. When visual contact is prevented, male and female alternate their calls with a very regular answer delay. This alternation is not the same at short (Far No Visual) or long distance (Close No Visual). This alternation of calls disappeared when visual contact is allowed. Figure A2 shows an eExample of the cumulative number of calls emitted during one minute in each condition for a single pair. In the Baseline condition, we extracted this minute from a period when the call rate was high (a). When visual contact was prevented and only acoustic contact was possible (Far No Visual and Close No Visual, dotted lines), the cumulative number of calls of the male and the female were highly correlated, with periods where both partners were calling (arrow on A) and periods where both partners remained silent. In these conditions, graphs of call timing show that male and female alternated their calls (b). In the Far No Visual context the graph shows a regular alternation of partners call, whereas in the Close No Visual context the female answered the male almost systematically after a short delay. This alternation of calls disappeared when visual contact was allowed (Reunion and Baseline) and curves of cumulative number of calls were no longer symmetrical. Call timings were also different, with more variable delays of response to the partner.

914 FIGURES

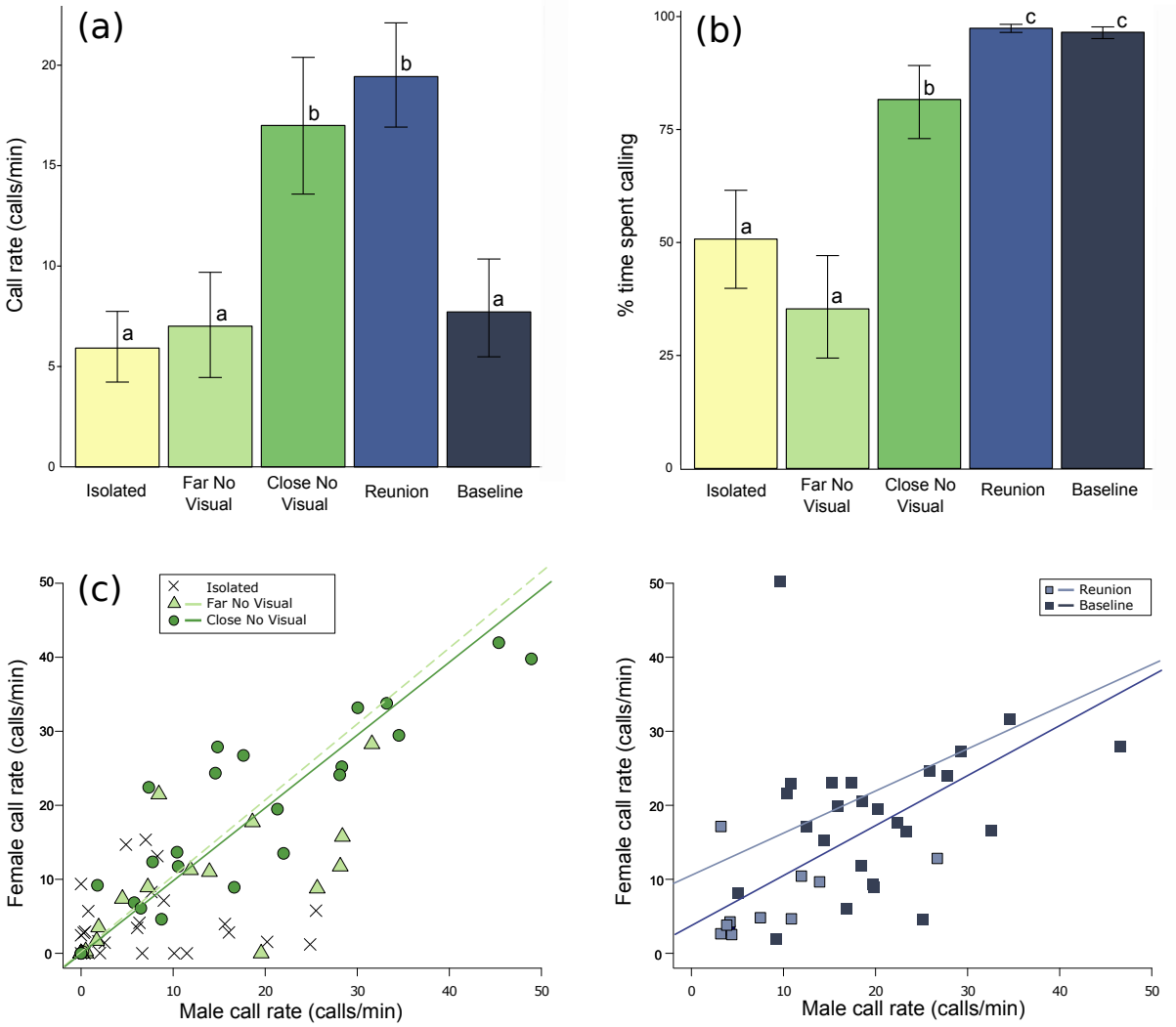
915 Figure 1



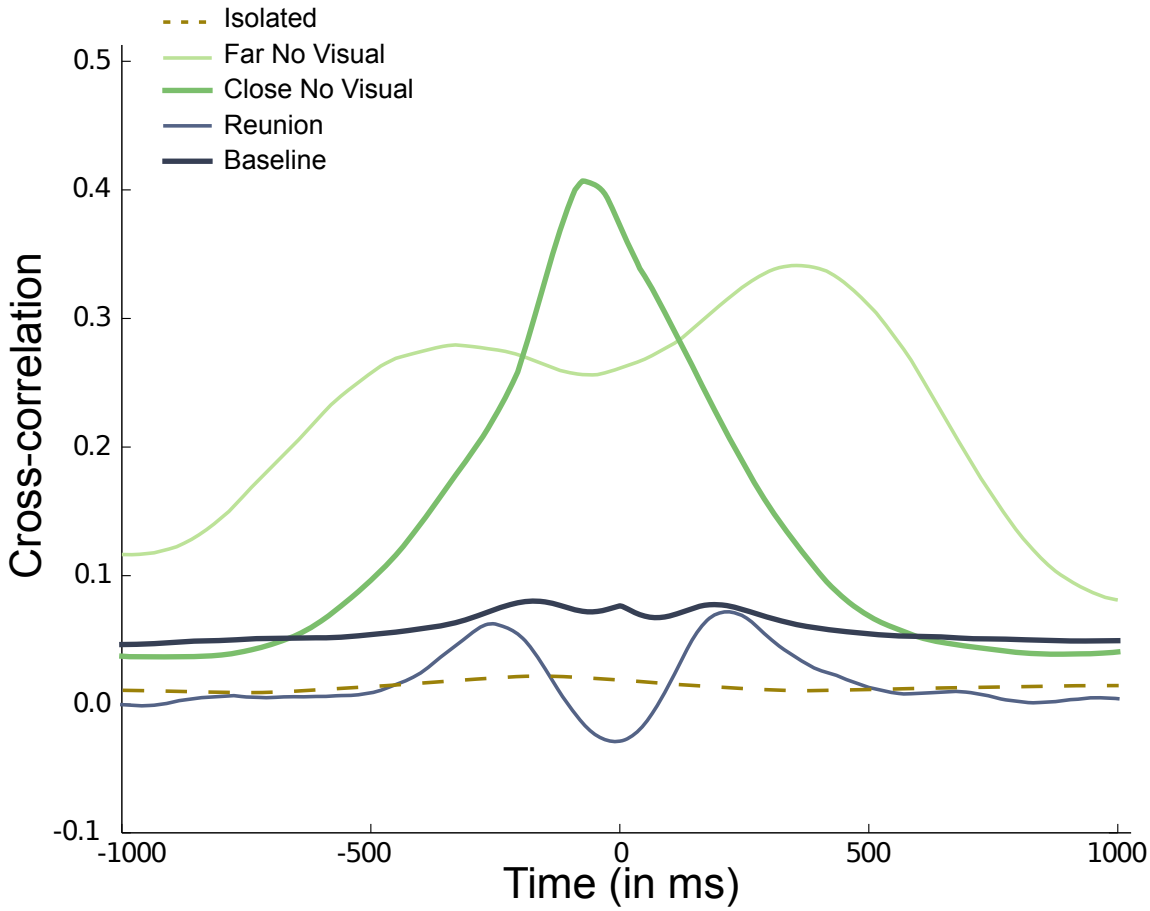
916

917

Figure 2:


(a)
Call rate (calls/min)
0
5
10
15
20
a
a
b
b
a
Isolated
Far No Visual
Close No Visual
Reunion
Baseline
(b)
% time spent calling
0
25
50
75
100
a
a
b
c
c
Isolated
Far No Visual
Close No Visual
Reunion
Baseline
(c)
Isolated
Far No Visual
Close No Visual
Female call rate (calls/min)
Male call rate (calls/min)
0
10
20
30
40
50
Reunion
Baseline
Female call rate (calls/min)
Male call rate (calls/min)

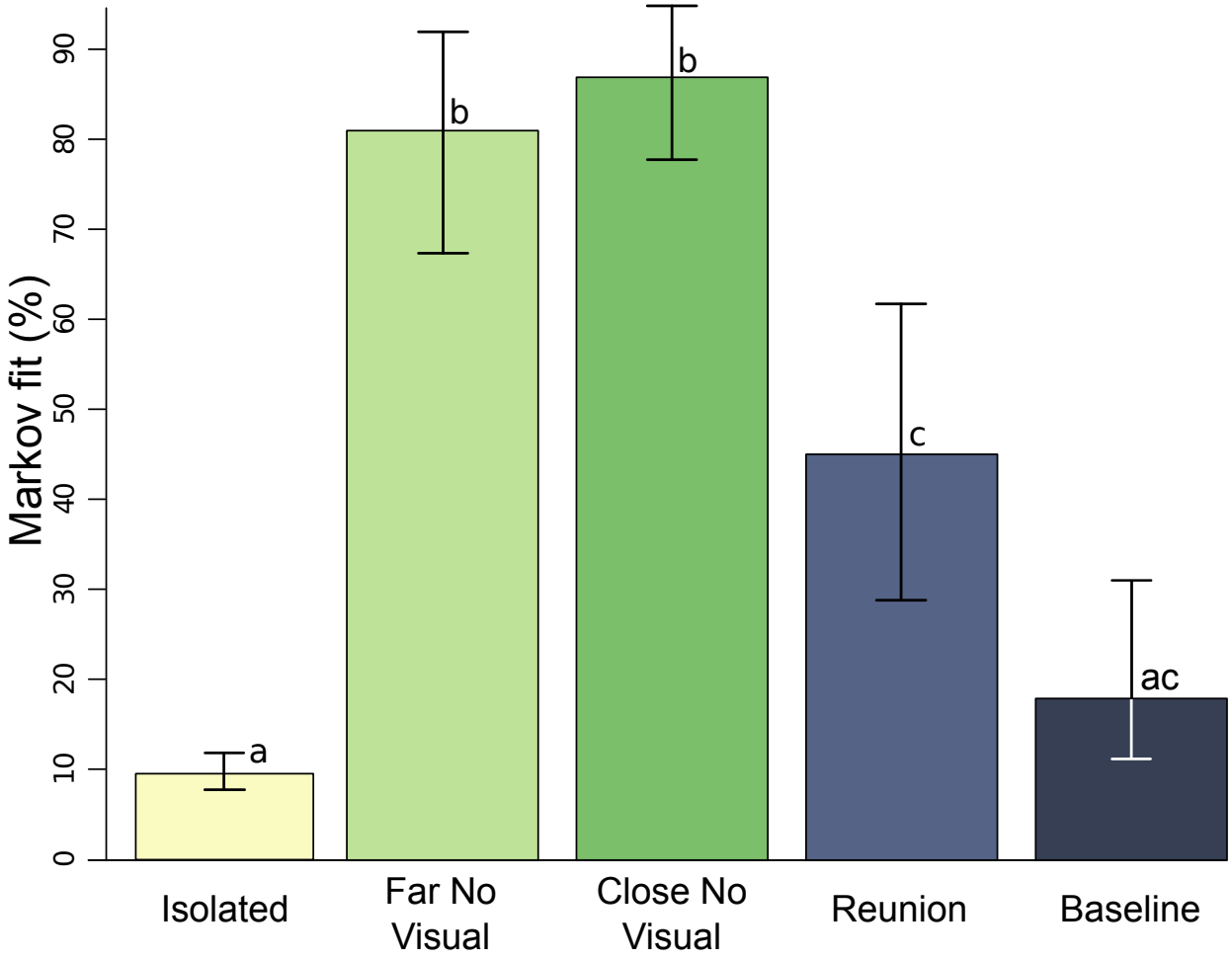
922 Figure 3:



923

924

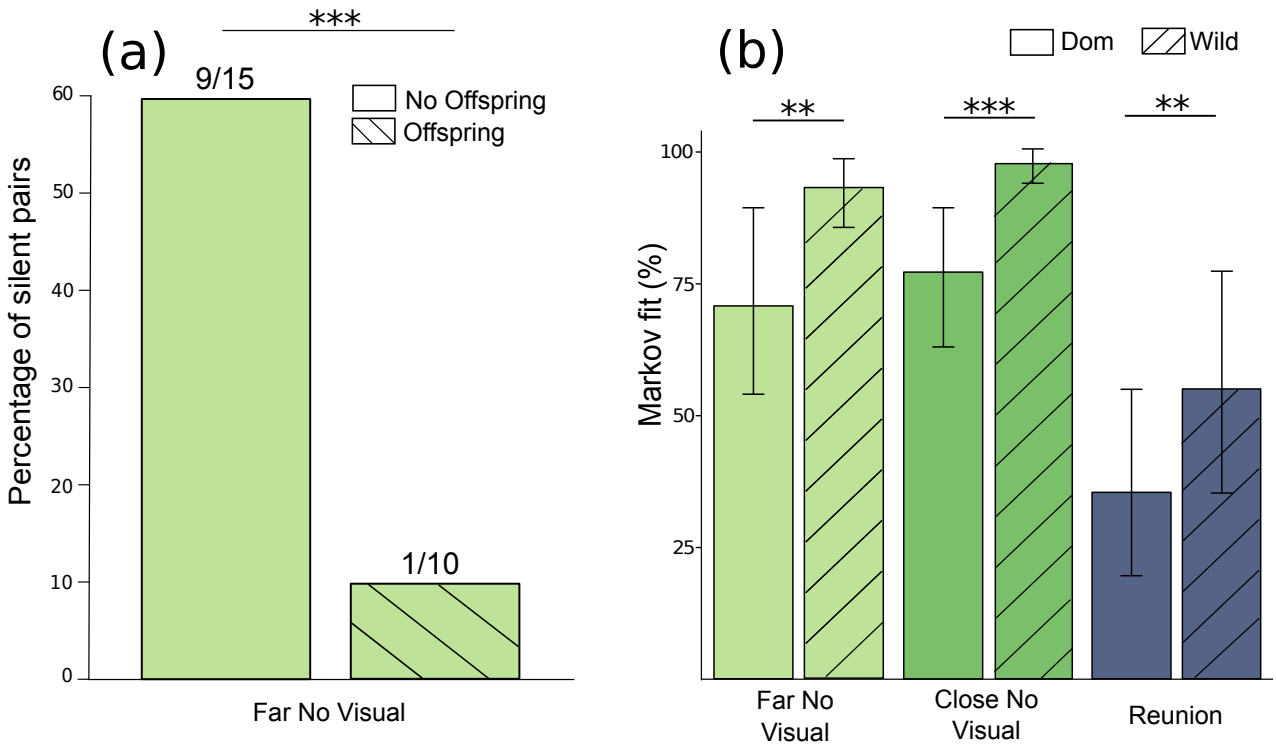
925 Figure 4:



926

927

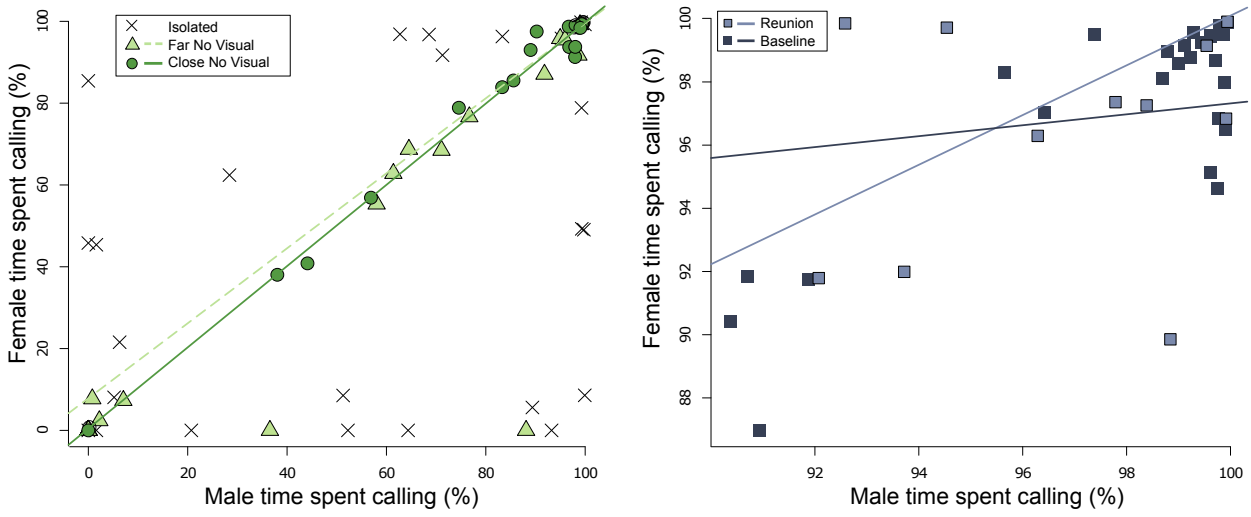
928 Figure 5:



929

930

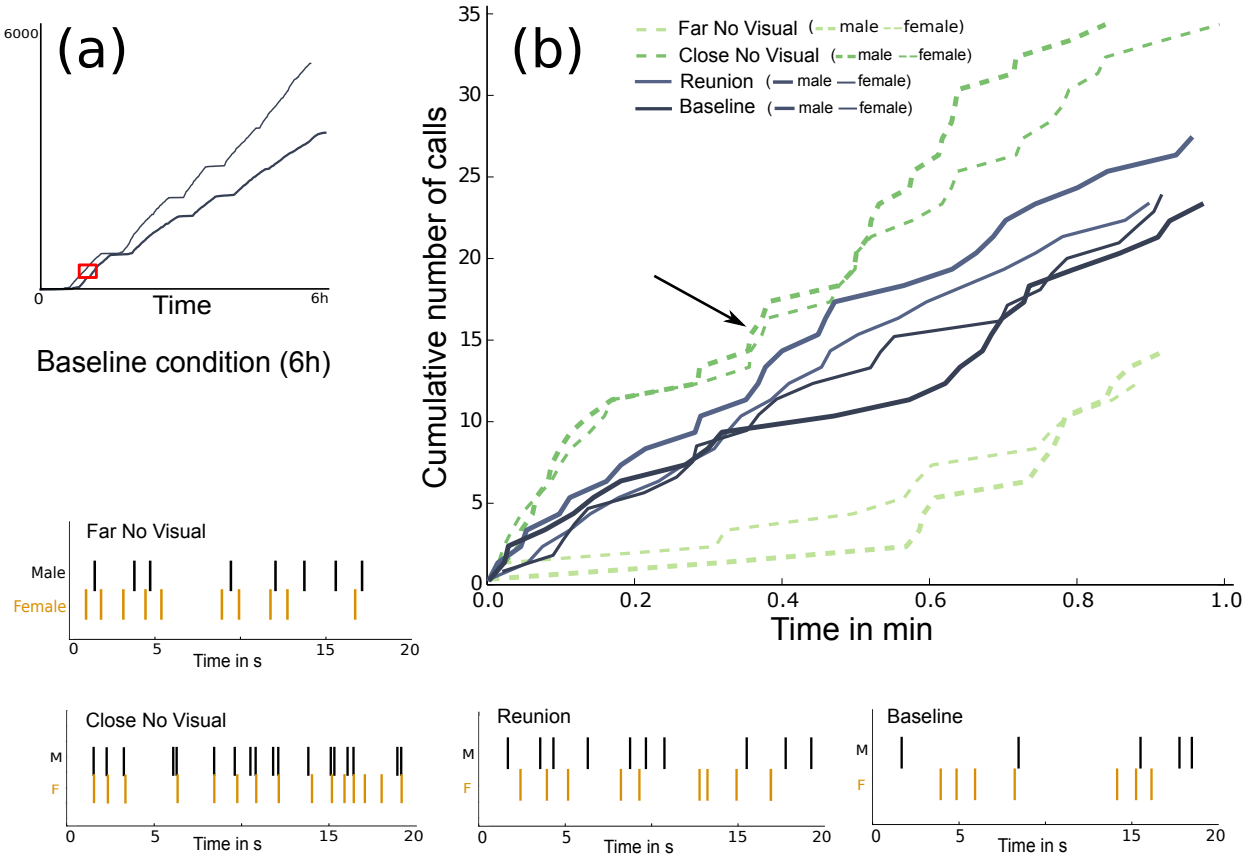
931 Figure A1:



932

933

Figure A2:


(a)
6000
0
Time
6h
Baseline condition (6h)
(b)
Far No Visual (male female)
Close No Visual (male female)
Reunion (male female)
Baseline (male female)
Cumulative number of calls
Time in min
Far No Visual
Male
Female
Time in s
Close No Visual
Reunion
Baseline
M
F

938 TABLES

939
940 Table 1: P-values of statistical tests on differences in call rate and time spent calling between

941 conditions (Isolated, Far No Visual, Close No Visual, Reunion and Baseline)

942

	Isolated	Far No Visual	Close No Visual	Reunion	Baseline
Call rate					
Isolated					
Far No Visual	$P = 0.87$ (1)				
Close No Visual	$P < 0.001$ (1)	$P = 0.001$ (1)			
Reunion	$P < 0.001$ (1)	$P < 0.001$ (1)	$P = 0.59$ (1)		
Baseline	$t_{34} = -1.06$, $P = 0.29$ (2)	$t = -0.24$, $P = 0.8$ (4)	$t_{33.8} = 3.03$, $P = 0.005$ (3)	$t_{34} = 4.66$, $P < 0.001$ (2)	
Time spent calling					
Isolated					
Far No Visual	$P = 0.15$ (1)				
Close No Visual	$P < 0.001$ (1)	$P < 0.001$ (1)			
Reunion	$P < 0.001$ (1)	$P < 0.001$ (1)	$P = 0.001$ (1)		
Baseline	$t_{24.6} = 6.81$, $P < 0.001$ (3)	$W = 18$, $P < 0.001$ (4)	$t_{24.8} = 2.45$, $P = 0.02$ (3)	$t_{34} = -0.8$, $P = 0.42$ (2)	

943

944 Call rate: (1) Pairwise Wilcoxon signed rank test after Friedman test: $\chi^2_3 = 37.75$, $N = 25$ pairs, $P <$

945 0.001 for global difference between conditions, (2) Student test, $N = 36$ pairs, (3) Student test with

946 Welch correction, $N = 36$ pairs, (4) Student test with permutations, $N = 36$ pairs; Time spent calling:

947 (1) Pairwise Wilcoxon signed rank test after Friedman test: $\chi^2_3 = 43.41$, $P < 0.001$ for global

948 difference between conditions, $N = 25$ pairs, (2) Student test, $N = 36$ pairs, (3) Student test with

949 Welch correction, $N = 36$ pairs, (4) Wilcox exact test (rank with ties), $N = 36$ pairs

950
951

952

953 Table 2: Differences in call rate and time spent calling between sexes

954

	Isolated	Far No Visual	Close No Visual	Reunion	Baseline
call rate	t=-1.79, N=50 P=0.071 (1)	t=-0.8, N=50 P=0.41 (1)	t_{48}=-0.22, N=50 P=0.82 (2)	t_{48}=-0.46, N=50 P=0.64 (2)	t=-0.62, N=22 P=0.58 (1)
time spent calling	t=-0.86, N=50 P=0.37 (1)	t=-0.44, N=50 P=0.64 (1)	t=-0.013, N=50 P=0.99 (1)	t=-0.78, N=50 P=0.45 (1)	t_{20}=-0.23, N=22 P=0.82 (2)

955

956 P-values of statistical tests in each condition (Isolated, Far No Visual, Close No Visual, Reunion

957 and Baseline). (1) Student test with permutations, (2) Student test

958

959 Table 3: Correlation coefficients (R^2) between males and females for call rate and time spent
960 calling
961

	Isolated	Far No Visual	Close No Visual	Reunion	Baseline
call rate	t_{23}=-0.11, P=0.91	t_{23}=5.57, P<0.001	t_{23}=9.64, P<0.001	t_{23}=1.03 P=0.31	t_9=1.55, P=0.15
time spent calling	t_{23}=-1.79, P=0.086	t_{23}=0.61, P<0.001	t_{23}=53.0, P<0.001	t_{23}=7.28, P<0.001	t_9=0.63, P=0.54

962

963 Correlation coefficients and *P*-values are given for the five conditions (Pearson's correlation test).

964

Table 4: Differences in Markov fit between conditions

	Isolated	Far No Visual	Close No Visual	Reunion	Baseline
Isolated					
Far No Visual	$z = 11.23$ $P < 0.001$				
Close No Visual	$z = 14.2$ $P < 0.001$	$z = 1.59$ $P = 0.483$			
Reunion	$z = 7.74$ $P < 0.001$	$z = -6.02$ $P < 0.001$	$z = -8.99$ $P < 0.001$		
Baseline	$z = 1.55$ $P = 0.512$	$z = -4.78$ $P < 0.001$	$z = 5.85$ $P < 0.001$	$z = 1.90$ $P = 0.297$	

P-values of generalized linear model with binomial family are given for each condition (Isolated, Far No Visual, Close No Visual, Reunion and Baseline).

Table A1: Detailed model results for the correlation between male and female vocal activity.

female call rate ~ male call rate * condition + 1|pair

Fixed effects:

	Value	SE	df	t-value	p-value
(Intercept)	4.175908	1.857622	68	2.247985	0.028
male call rate	0.017625	0.172131	68	0.102390	0.919
condition Far No Visual	-2.898846	2.348175	68	1.234510	0.221
condition Close No Visual	0.167192	2.590470	68	0.064541	0.949
condition Reunion	9.401256	3.482469	68	2.699595	0.009
condition Baseline	0.048052	3.713571	33	0.012940	0.990
male call rate : condition Far No Visual	0.554648	0.207746	68	2.669837	0.009
male call rate : condition Close No Visual	0.770939	1.189509	68	4.068087	<0.001
male call rate : condition Reunion	0.242596	0.213680	68	0.213680	0.260
male call rate : condition Baseline	0.295601	0.340945	33	0.867005	0.392

female time spent calling ~male time spent calling * condition + 1|pair

Fixed effects:

	Value	SE	df	t-value	p-value
(Intercept)	26.06386	7.17033	68	3.634962	0.001
male time spent calling (TSC)	0.35428	0.10495	68	3.375780	0.001
condition Far No Visual	-26.38714	9.13475	68	2.888655	0.005
condition Close No Visual	-25.83110	14.16946	68	1.823012	0.073
condition Reunion	-15.66287	128.87304	68	0.121537	0.904
condition Baseline	46.39632	212.17869	33	0.218666	0.828
male TSC : condition Far No Visual	0.51898	0.14607	68	3.552836	0.001
male TSC : condition Close No Visual	0.64426	0.17569	68	3.667123	0.001
male TSC : condition Reunion	0.53188	1.31955	68	0.403077	0.688
male TSC : condition Baseline	-0.10707	2.19468	33	0.048785	0.961

Detailed results are shown for the call rate and time spent calling (111 observations on 36 pairs). Call rate random effects standard deviation: intercept=2.74, residual=6.06. Time spent calling random effects standard deviation: intercept=0.002, residual=20.75

Table A2: Detailed model results for the probability of emitting at least one call.

probaOneCall ~ condition * Offspring + 1|pair

Fixed effects:

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	1.8877	0.8498	2.221	0.026
condition Far No Visual	0.9544	1.0095	0.945	0.344
condition Close No Visual	0.9545	1.0095	0.945	0.344
No Offspring	0.7339	1.0995	0.668	0.504
condition Far No Visual : No Offspring	-4.5790	1.4243	-3.215	0.001
condition Close No Visual : No Offspring	-0.9545	1.3181	-0.724	0.469

Detailed results are shown for the Far No Visual and the Close No Visual conditions (78 observations on 25 pairs). Random effects (pair) variance=2.24, standard deviation=1.50.

Table A3: Detailed model results for the Markov fit of calling sequences.

markov fit ~ condition + 1|pair

Fixed effects:

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	-2.2927	0.3308	-6.931	< 0.001
condition Far No Visual	4.0761	0.3628	11.235	< 0.001
condition Close No Visual	4.6465	0.3267	14.221	< 0.001
condition Reunion	2.1112	0.2729	7.736	< 0.001
condition Baseline	0.9677	0.6260	1.546	0.122

markov fit ~ condition * Type + 1|pair

Fixed effects:

	Estimate	SE	z value	Pr(>|z|)
(Intercept)	-2.2795	0.5123	-4.450	<0.001
condition Far No Visual	3.3426	0.5061	6.604	<0.001
condition Close No Visual	3.7069	0.4648	7.975	<0.001
condition Reunion	1.5910	0.4442	3.581	<0.001
Type Wild	-0.3842	0.8030	-0.478	0.632
condition Far No Visual : Type Wild	2.2228	0.9593	2.317	0.020
condition Close No Visual : Type Wild	3.2109	0.9654	3.326	<0.001
condition Reunion : Type Wild	1.3740	0.7335	1.873	0.061

For the first model, detailed results are shown for all conditions (107 observations on 32 pairs), random effects (pair) variance=1.40, standard deviation=1.18. For the second model, detailed results are shown for the Isolated, Far No Visual, Close No Visual and Reunion conditions (78 observations on 25 pairs), random effects (pair) variance=1.33, standard deviation=1.15.